

# Assessment as Inquiry

GINETTE DELANDSHERE

*Indiana University*

*For more than 10 years now, arguments have been constructed regarding the need for new forms of educational assessment, and for a paradigm shift with a focus on supporting learning rather than on sorting and selecting students. The call for change in assessment follows an almost unanimous recognition of the limitations of current measurement theory and practice. The conceptions of learning represented by theories of learning and cognition appear strikingly different from those implied in current educational assessment and measurement practices. Indeed, most educational measurement specialists are still working from century-old understandings and behaviorist perspectives. Although the call for change is clear, the proposals and recommendations being put forward have limitations of their own and are unlikely to yield the kinds of fundamental changes envisioned by researchers. These limitations lie either in the focus of the work, in the lack of a clear articulation of the theories and concepts, in the nature of the assumptions made about learning (many of which remain implicit and unchanged), in the exclusion of certain conceptions of learning, or in some combination of these problems. This article explores the possibility of using inquiry as a way to understand, and hence to assess, learning. After an initial review of the assessment literature in which the need for change has been asserted and analysis of the theoretical and epistemological foundations that seem to undergird these writings, the focus shifts to the meaning of learning, knowing, and teaching implied in this literature and to the limitations of its recommendations. Later sections consider notions of learning that seem to be excluded from current assessment practices and begin to uncover similarities between learning, knowing, and inquiring that could make inquiry an appropriate metaphor for what we currently know as educational assessment. Finally, there is discussion of important issues that would need to be considered in an inquiry framework for assessment.*

For more than 10 years now, arguments have been constructed regarding the need for new forms of educational assessment (e.g., Shepard, 1989, 2000), for a paradigm shift with a focus on supporting learning rather than on sorting and selecting students (Gipps, 1994, 1999; Resnick & Resnick, 1992; Shepard, 2000). Some researchers have pointed to the dangers of trends, such as the minimum competency movement in the 1970s and to the negative consequences of state assessment practices for classroom assessment and for the curriculum (e.g., Delandshere & Jones, 1999; Madaus,

1994; Madaus & McDonagh, 1979; Shepard, 1989, 2000). Some have questioned measurement assumptions in the context of complex forms of assessment (e.g., Delandshere, 2001; Delandshere & Petrosky, 1998), and others (e.g., Embretson, 1993, 1998; Frederiksen, Mislevy, & Bejar, 1993) have considered the development of a new test theory grounded specifically in more recent developments in cognitive theory. Although all of these efforts address the problems and limitations of today's educational measurement and assessment practices, working with existing concepts, methods, and technology toward improving the current system, they seem to be buried under the renewed call for accountability and for increased measurement of learning outcomes as a way to evaluate the efficiency of the educational system. The sense of urgency and the pressure to use current forms of assessment for political aims are compromising the search for new ways to study and assess learning and to make claims about the state of education in this country.

Currently most state assessment programs and many classroom assessment practices<sup>1</sup> resemble poorly designed research projects for which the most fundamental questions have not been asked. "What do students know?" seems to have been one of the main assessment questions. But the question "What does it mean to know?" is rarely asked, let alone discussed or answered, other than implicitly through the content of the questions included on the test. Such a question is critical, and yet as Gill (1993) notes, "Among the many and various articles and books on the quality and direction of American education, one searches in vain for an in-depth discussion of how knowing takes place, of who knowers are, and of what can be known" (p. 1). Until we come to grips with, or at least frame the issue of, knowledge and knowing in ways that can guide education practices (including assessment), the enterprise of education runs the risk of being fruitless and counterproductive. In its current state, assessment appears to be a process of collecting data about phenomena or constructs that we have not adequately defined, to answer questions that we have not articulated, and on the basis of which we draw inferences about the quality of the education system. Such practice has the allure of naïve inductive empiricism, which has been seriously criticized in most areas of social science research for quite some time (e.g., Cronbach & Meehl, 1955; Douglas, 1971; Feyerabend, 1978; Lakatos, 1970; Polanyi, 1958; Wallace, 1971). Before we can generate valid inferences about the important project of education, we need to reconnect our educational practices to theoretical and philosophical considerations to clarify the assumptions we make about learning and teaching.

In this article I explore the possibility of using inquiry as a way to understand, and hence to assess, learning. I first review the assessment literature in which the need for change has been asserted and analyze the theoretical and epistemological foundations that seem to undergird these

writings. I focus primarily on the meaning of learning, knowing, and teaching implied in this literature and consider the limitations of its recommendations. I then consider notions of learning that seem to be excluded from current assessment practices and begin to uncover similarities between learning, knowing, and inquiring that could make inquiry an appropriate metaphor for what we currently know as educational assessment. Finally, I discuss important issues that would need to be considered in an inquiry framework for assessment.

### CALLS FOR CHANGE IN ASSESSMENT

The conceptions of learning represented by theories of learning and cognition appear strikingly different from those implied in current educational assessment and measurement practices. Most educational measurement specialists have been working from century-old understandings and behaviorist perspectives (Mislevy, 1993; Shepard, 1991, 2000), according to which *to learn* means to accumulate knowledge, and learning results from teaching that is organized to impart that knowledge in an atomized, sequential, and hierarchical manner (Skinner, 1954). The task of assessment is then to observe whether individuals can reproduce this knowledge when they are stimulated to do so. This behaviorist conception of knowledge is articulated in simple, essentialist, and mechanistic terms and is well suited for the application of measurement axioms—based on estimating the probability of recurrence of correct responses given a particular stimulus, frequently a test question. This conception of assessment still underlies most practice of assessment today, whether in the form of end-of-unit tests, end-of-semester high school examinations, state- and district-mandated tests, college entrance examinations, and so on. In this perspective, knowledge is abstracted and generalized, decontextualized and defined independently of the individual's activity of learning. This conceptualization is usually implied, however, as the assessment and measurement literature has not traditionally addressed these issues explicitly. For example, one is hard-pressed to find explicit definitions of learning or knowledge in measurement and assessment textbooks (e.g., Allen & Yen, 1979; Crocker & Algina, 1986; Hopkins, Stanley, & Hopkins, 1990; Popham, 1990) or in more technical writings (see, for example, the *Journal of Educational Measurement*), as if these definitions were well understood and agreed on.

Researchers have attempted to address this conceptual vacuum in the assessment literature, calling for the integration of new understandings of human cognition (Mislevy, 1993; Snow & Lohman, 1989; Embretson, 1993), for theory- and construct-driven assessment (Messick, 1994), for changing the nature of assessment practices, and for integrating assessment with instructional practice (Gipps, 1994, 1999; Resnick & Resnick, 1992; Shepard,

2000; Wiggins, 1993). In 1993, for example, Frederiksen et al. (1993) edited a collection of essays proposing new developments in test theory. At issue was the recognition that educational measurement was facing a crisis: "The essential problem is that the view of human abilities implicit in standard test theory . . . is incompatible with the view rapidly emerging from cognitive and educational psychology" (Mislevy, 1993, p. 19). In response, many of the chapters in the book call for the integration of new statistical models with new developments in cognitive psychology, with these two strands of research considered as the foundation of a new test theory. Existing psychometric theories are analyzed in terms of their limitations. Their restrictive assumptions (e.g., constant universe score across observations) are viewed as incompatible with those of cognitive theory (Lohman & Ippel, 1993) and with the substantive theories that define the complex ability and task structures (Haertel & Wiley, 1993). Included here is a serious criticism of current testing practices and of the right-wrong scoring schemes they typically employ. Even the most promising applications of contemporary test theories are regarded as inadequate. As Embretson (1993) states: "Although the item response theory models that are typically applied have many advantages over earlier testing methods, they have little connection to concerns of cognitive theory about the processes, strategies, and knowledge structures that underlie item solving" (p. 125). She then proposes that psychometric models be developed that will allow for the direct integration of "cognitive processing variables" (p. 126).

In their review of learning and cognition research, Pellegrino, Baxter, and Glaser (1999) reaffirm the calls for similar changes in assessment design and practice to be better integrated with learning and instruction. Their argument follows Glaser's (1976) initial aim to develop a psychology of instruction that could be used prescriptively to match learners with optimal instructional treatments. In a brief retrospective, they conclude that the aptitude-treatment interaction (ATI) studies of the 1960s, although promising, yielded many contradictory results because of insufficient theoretical work and inadequate aptitude measures. Similarly, they claim that the cognitive process analyses of aptitude and achievement used to study individual differences in the 1970s and 1980s were not particularly instructionally relevant. They seem to regard the study of expertise and competence that emerged in the 1980s and 1990s as most useful for instructional purposes, describing it as an approach "that [is] no longer linked to psychometrically defined constructs of aptitude or intelligence" (p. 317), although they later acknowledge that thus far no approach has "fully confront[ed] the conceptual conflict between theories of cognition and learning and contemporary models and methods of psychometric techniques" (p. 333). To do so, they state, "will require the integration of theories of knowledge and instruction with new psychometric models that describe acquired competence in subject-

matter learning” (p. 338). The new test theory should be informed by the study of competence and expertise and will “account for differences among and within individuals in terms of knowledge structures such as schemas, mental models, and semantic and procedural networks” (p. 344).

To integrate advances in cognitive psychology and in technology, Mislevy and his colleagues (2001) propose that understandings of “the nature of knowledge in [a] domain of interest, how students learn it and how they use their knowledge” (p. 3) be combined with understandings of “what and how much we learn about students’ knowledge from what they say and do” (p. 4). They suggest that three complex models should be defined and coordinated for the assessment to be coherent: (1) the student model, which defines what should be assessed; (2) the evidence model, which lays out the behaviors that constitute evidence of the constructs included in the student model; and (3) the task model, which “includes specifications for the environment in which the student will say, do, or produce” (p. 18) the targeted behaviors, and for the work produced. With this approach, they work from a theory of performance (the student model) in a particular domain, define the types of evidence (the evidence model) that are needed consistent with the theory, and map these against task specifications (the task model) that will elicit the needed evidence.

The case for a need to change assessment practices has been made and seems to converge on the integration of theoretical considerations about learning and cognition and on making assessment an integral part of learning and instruction. In the next section, I analyze the limitations of these proposals for change by examining some of the ambiguities and omissions that may prevent fundamental changes in our educational practices and our understanding and assessment of learning.

### LIMITATIONS OF THE PROPOSALS FOR CHANGE

Although the call for change is clear, I contend that the proposals and recommendations being put forward are unlikely to yield the kinds of fundamental changes envisioned by researchers. Their limitations lie either in the focus of the work, in the lack of a clear articulation of the theories and concepts, in the nature of the assumptions made about learning (many of which remain implicit and unchanged), in the exclusion of certain conceptions of learning, or in some combination of these problems.

### FOCUS

Researchers often focus their work either on classroom assessment or on large-scale assessment used for accountability purposes, as if assessment could mean something different in these two contexts. As long as this

divide continues, the status quo will prevail. The way we assess learning is so closely related to the kind of learning we value and how we conceive of it that it does not seem possible that two completely different understandings could coexist without one overtaking the other. The political power of accountability testing and its effect on classroom assessment practices make it impossible to change one without changing the others. Current arguments for new forms of assessment often fail to address this limitation. Shepard (2000), for example, recognizes "the pervasive negative effects of accountability tests and the extent to which externally imposed testing programs prevent and drive out thoughtful classroom practices" (p. 9) but focuses her attention exclusively on classroom assessment. In effect, positing differences between "the kind of assessments used to give grades or to satisfy the accountability demands of an external authority . . . [and] the kind of assessment that can be used as a part of instruction to support and enhance learning" (p. 4) recognizes their possible coexistence. Similarly, Gipps (1999) in her review of sociocultural aspects of assessment recognizes that "many of the new developments in assessment are designed for use by teachers at the classroom level" (p. 375), which she also sees as coexisting with the more traditional forms of assessment. Referring to the British experience, she believes that "It is possible to design an assessment program with different features and purposes at different levels of the school system" (p. 384). The question remains: Can assessment whose function is to support learning coexist with assessment whose function is to reward or penalize learners, teachers, or schools? And why are we calling these two very different processes by the same name?

On the other hand, because of their complexity, some of the newly developed assessment models, tasks, and psychometric models (e.g., Embretson, 1998, 1993; Haertel & Wiley, 1993; Lohman & Ippel, 1993; Mislevy et al., 2001) seem quite removed from the classroom—the context in which students learn—and are evidently inaccessible to teachers. The models proposed by Mislevy and his colleagues (2001), for example, require elaborate analysis of each specific domain and task of interest and also involve complex statistical modeling that makes use of "classical test theory, item response theory, latent class models and factor analysis" (p. 14). Such models might be considered experimental or for use in external assessment, but they are unlikely to be used by teachers in the classroom.

From reading these different lines of research, one wonders whether there is some common ground, or whether researchers in these areas are moving in very different directions. It also appears—although it is not at all clear, as we will see later—that classroom assessment tends to be framed within the rhetoric of social constructivist and situative learning theories with a primary purpose of supporting learning (Gipps, 1994, 1999). Meanwhile, most of the external assessment literature is grounded in cognitive

psychology and psychological constructivism<sup>2</sup> (e.g., Piaget). These latter theoretical perspectives seem to accommodate different assessment purposes (e.g., assessing readiness to learn or assigning grades), and their underlying assumptions, as we will see later, are compatible with the application of psychometric models.

The important point here, though, is that if different people with different orientations address classroom assessment and external assessment separately, there is only a remote chance for change in educational practices. The renewed push for the use of multiple-choice tests for accountability purposes (e.g., President G. W. Bush's proposal for yearly testing in mathematics and reading) will further inhibit changes at the classroom level.

### THEORETICAL AND CONCEPTUAL ARTICULATIONS

Researchers calling for changes in assessment are often not clear about or fail to specify the theoretical perspectives and the definitions of learning and knowing from which they work. In the Frederiksen et al. (1993) volume, the authors seem generally to be working from a cognitivist perspective—a broad category of theories that explain how the mind processes information. Memorization, schemas, mental processes, and task and ability structures are key concepts for information processing. In this perspective, learning is not defined consistently (or sometimes, explicitly), other than through the mental activities that presumably make it occur, and knowledge seems to be equated with factual information. At times “learning is . . . seen as a progression through qualitatively different models in a domain—as successive approximation to a target domain” (Snow & Lohman, 1993, p. 11). At other times learning is implicitly defined through knowledge and understanding in an acquisitive-deficiency framework, with learning deficiencies being categorized with such labels as “wrong knowledge,” “inert knowledge,” and “lost knowledge” (Feltovich, Spiro, & Coulson, 1993). The concept of aptitude is also used and defined within a hierarchical structure with cognitive processing. Embretson (1993) states: “Aptitude results from solving complex tasks that require correct outcomes from several stages of the task” (p. 127). The meaning of knowing can be inferred here only in relation to the complexity of the tasks that one can solve.

Pellegrino et al. (1999) work within a similar information-processing framework; they use categories and concepts that emerged from early empirical work in measurement without redefining them. Aptitude, ability, achievement, competence, performance, and proficiency are used throughout as if these concepts were universally meaningful and unchallenged. The authors distinguish between cognitive aptitude and academic achievement (p. 309) and state that “intelligence and aptitude tests are essentially measures of scholastic ability” (p. 313). Expertise is “the continuous acquisition and



restructuring of domain-based knowledge" (p. 317), and "competence is defined, in part, by the extent to which knowledge and skills are transferable" (p. 335). But one is left wondering about the theoretical connections and implications of these concepts in light of other work where the distinction between achievement and ability, for example, is being questioned (Sternberg, 1998). For Pellegrino et al. (1999), knowledge and competence are understood as things that are acquired and transferred, and learning is equated with cognition, whose configuration, structure, complexity, and processing can be uncovered and used prescriptively to design better forms of assessment and guide instructional practices. In this perspective, what it means to learn or to know (the focus of educational assessment) is not entirely clear. At times knowledge simply means information, but in other instances it also seems to refer to cognitive structure.

From reading this work, one gets the impression that if we just had the "right" theories of cognition, better task analyses, and the technology to model them, standards and criteria for competent performance could be specified. The problem of assessing and improving achievement would then be solved. In this perspective, cognition is located within the individual, and "assessments can be designed with predictable cognitive demands" (Pellegrino et al., 1999, p. 339). Paradoxically, however, knowledge appears to be definable outside the individual and independently of context (i.e., heavy reliance on the notion of transfer from context to context and models of expertise for expected performances). In this perspective, the goal is to articulate substantive theories of performance to understand the

common developmental trajectories through which learning progresses. . . . From theory and data, one [then] posits probabilities for the ways that students with different configurations of knowledge, skills and other distinguishing characteristics of performance (e.g. problem representation, strategy use) will solve problems, answer questions and so on. (p. 340)

Although this literature raises important theoretical concerns, it seems bound by the same categories and ideas that emerged from a quantitative and acquisitive view of knowing and learning and the predictive, deterministic psychometric framework.

Constructivist perspectives of learning are also acknowledged when Masters and Mislevy (1993) state that "learning is increasingly being recognized as an active process through which students construct their own interpretations, approaches, and ways of viewing phenomena, and through which learners relate new information to their existing knowledge and understandings" (p. 220). In the discussion of how to assess this constructed knowledge, however, there also seems to be an exclusive recourse to cognitive psychology and complex task analysis, as if it would be possible to



anticipate and predict the cognitive structures, processes, and skills involved in all individual constructions. More puzzling is how cognitivist, constructivist, and situativist views of learning have been lumped together and contrasted to behaviorist perspectives, without much regard for the different assumptions implicit in these views about learning. The following view has shaped the work of Mislevy and his colleagues (2001):

The cognitive perspective includes both the constructivist tradition originated by Piaget and the information-processing tradition developed by Newell and Simon, Chomsky, and others. The focus is on patterns and procedures individuals use to acquire knowledge and put it to work. The situative perspective focuses on the ways individuals interact with other people in social and technological systems, so that learning includes becoming attuned to the constraints and affordances of these systems. In this paper, we use the term “cognitive psychology” broadly to encompass both of these perspectives. (p. 22)

Such categorization fails to recognize the many variants of constructivism, from the radical psychological constructivist to the radical social constructionist (McCarty & Schwandt, 2000; Phillips, 2000), and the different implications that these views would have for assessment. Further, the work of Mislevy et al. (2001) is grounded primarily in cognitive psychology, and it is unclear how it rests on any constructivist or situative perspectives. A similar eclectic categorization appears in Shepard (2000) when she elaborates “the principles of a ‘social-constructivist’ conceptual framework, borrowing from cognitive, constructivist, and sociocultural theories” (p. 6). Although she acknowledges the existence of these different perspectives, “which are at times warring with each other” (p. 6), most of the strategies that she recommends for assessment (e.g., assessment of prior knowledge, feedback, transfer, self-assessment) find their roots in cognitivist traditions.

There is a sense that the field is reacting to its initial inductive empiricism and attempting to move toward a more deductive theoretical process. The danger here lies in treating theories as truths rather than as frames for temporary understanding. In so doing, we run the risk that we will simply fix learning in different categories without really furthering our understanding. The theories of learning and cognition now being considered are more complex and the statistical models more sophisticated, but are these new developments really working from different assumptions with different implications for assessment?

#### ASSUMPTIONS ABOUT LEARNING AND KNOWING

The call for change in assessment is in part a call to move away from simple, mechanistic behaviorist notions of learning toward cognitive, con-

structivist, and situative representations. As we have seen, researchers tend to collapse these last three theoretical orientations while working primarily within a cognitivist perspective. It seems, however, that behaviorist and cognitive theories make similar assumptions about human behaviors. Cognitive theory is often thought to be fundamentally different from behaviorism because in the former the mind is recognized as playing a role in the learning process or the processing of information, and the models of memory it entails are often regarded as qualitatively different from the behaviorist explanations of learning. But as Slife and Williams (1995) point out, both perspectives are grounded in the same epistemology and work from the same ideas rooted in empiricism—"the notion that our learning and memory are primarily derived from our experience of events of the world" (p. 67). Referring to Solso's work (1991) in cognitive psychology, they state:

Sensory experiences are symbolized and processed in other ways [than according to behaviorist explanations], but such experiences are nevertheless the root of knowledge, because all knowledge is 'derived' from these experiences. . . . Nearly all cognitive models of knowing can be boiled down to this basic formula: INPUT-PROCESSING-OUTPUT. The environmental input with which the formula begins is processed in the mind and then output in the form of behavior. This type of theorizing is clearly within the tradition of empiricism. (p. 69)

This basic cognitive model is a more complex representation of the stimulus-response model, but it rests on the same determinist assumption (Slife & Williams, 1995). The information (or input) received from the environment determines the cognitive processing and the person's response, just as stimuli and reinforcement produce human behaviors. There are different views about how the information is mediated and about the origins of the mental processing involved, but fundamentally it seems that "All of the ideas and assumptions that are implicit in behaviorism are also embedded in cognitive theory" (p. 45).

Such determinism makes it possible to control people's learning. Behaviorist control operates mechanistically through the proper contingencies of stimuli and reinforcement to yield a particular response. Cognitive processing can be manipulated by structuring the information received by the individual in such a way as to require particular responses. This deterministic assumption has major implications for assessment. If it is possible to predict what people learn (given particular stimuli) or how they process information (structure it in a particular way), then it is also possible to design assessments using predetermined correct responses or predetermined processing paths (single or multiple) or models of performance. This seems to be the basis for developing new assessment models (Embretson, 1993, 1998; Frederiksen et al., 1993; Haertel & Wiley, 1993; Mislevy

et al., 2001). These are conceptually more sophisticated because of the theories of performance or ability from which they are derived, but they are still grounded in the same epistemological assumptions. Knowledge here is implicitly understood as information taken from the external world (Slife & Williams, 1995) that, if transformed and processed in appropriate ways, will produce specific responses. Such predictability (at least with some degree of certainty) is necessary for statistical modeling and makes cognitivist models of performance compatible with the psychometric tradition and its new developments.

The determinism and control that underlie cognitive theory also have implications for our understanding of the relationship between teaching and learning. If learning is controlled by the environment, then teaching consists of structuring the environment or information so as to produce the desired outcomes or responses. This assumption is at the core of many forms of accountability testing that have linked teaching and school quality to students' performance. Without this assumption, such accountability practices would be unjustifiable.

Other theoretical perspectives (constructivist and situative) that have been collapsed under the cognitive theory umbrella make different assumptions about learning and knowing. Piaget, for example, who is associated with what Phillips (2000) calls psychological constructivism, falls into the broad category of structuralist. His stage theory posits that human development is based on underlying stage structures that cannot be observed. These structures constitute the explanation for human behaviors; they are innate and universal and are grounded in rationalist epistemology, which assumes that knowledge results from logical thinking and reasoning (Slife & Williams, 1995). Such explanations are also greatly deterministic and have recently been challenged (Cannella, 1998; Zimiles, 2000).

Although it does not seem possible to bring order to the constructivist landscape, it is important to realize that these theorists are not presenting a unitary view of learning (Phillips, 1995, 2000). The various constructivist schools of thought have not been duly recognized in the assessment literature. Moreover, how their respective assumptions differ from each other and from cognitive theory has not been considered, nor have their implications for assessment. Many branches of constructivism and constructionism are viewed as postmodern theories (McCarty & Schwandt, 2000; Phillips, 2000; Slife & Williams, 1995) that evolved as a critique of empiricist epistemology and modernist assumptions.

One of the central ideas of this criticism is that contrary to what has been central to our intellectual tradition, there are no metaphysical absolutes, no fundamental and abstract truths, laws, or principles that determine what the world is like and what happens in it. . . . [Human]

behaviors will never be understood by looking for laws or principles . . . [they] can only be understood by 'reading' the broader context of life and history within which the behavior occurs. (Slife & Williams, 1995, pp. 54–55)

Language, in which knowledge is created, is accorded primary importance, and human actions and behaviors are regarded as texts.

If human behaviors cannot be predicted on the basis of a set of laws or principles, then they cannot be studied and assessed using the same methods and a priori models of performance that are currently used in assessment. The main critique of constructivism/constructionism is that in its extreme forms it is completely relativist. Because it is impossible to establish truth with any certainty, there is no basis for making judgments or assessing behaviors or their morality (McCarty & Schwandt, 2000; Slife & Williams, 1995). Indeed, for radical psychological constructivists (e.g., von Glasersfeld), learning occurs through individual constructions, which cannot be separated from the individual or from the environment because these do not exist independently. The external world cannot be known because it is simply an individual's construction. In this perspective,

learning is primarily a process of assimilation and accommodation to disturbances in the subject's experiential environment. . . . there is to be no notion of correct solution, no external standard of right and wrong. As long as a student's solution to a problem achieves a viable goal, it has to be credited. (McCarty & Schwandt, 2000, p. 49)

Social constructionists have developed equally radical views (e.g., Bergen) at the other end of the continuum, claiming that all human experiences are socially mediated and that our knowledge claims are simply linguistic constructions that are socially defined. Individuals cannot be conceived as such but only in terms of their linguistic relationships to others' constructions (Phillips, 2000; McCarty & Schwandt, 2000). From these radical perspectives, it is unclear what forms of assessment would constitute legitimate evidence of students' understanding or whether anything could count as such, but between these two extremes exist many variants (more or less radical or conservative) that are worth exploring. These indeed constitute new theories of learning and perspectives, which educators have overwhelmingly rallied around but for which assessment implications have not really been considered.

#### MISSING VOICES: SOCIAL CONSTRUCTIVIST PERSPECTIVES

There is a body of literature that includes, among others, historical, socio-cultural, and activity theories of learning. Here learning and knowing are

defined in ways that are quite different from those found in the assessment literature. Consider the following:

a theory of social practice emphasizes the relational dependency of agent and world, activity, meaning, cognition, learning and knowing. . . . learning, thinking, and knowing are relations among people in activity in, with, and arising from the socially and culturally structured world. . . . One way to think about learning is as the historical production, transformation, and change of persons. (Lave & Wenger, 1991, pp. 50–51)

A post-critical understanding of knowing begins with acknowledging (1) that there is no knowledge apart from active engagement between the knowing subject and that which is known, and (2) that knowing itself is a kind of doing, that knowledge is not had as much as it is done. In both of these respects participation plays a crucial role (p. 48). The key idea here is that knowledge is not a thing to be possessed but an activity to be engaged in. (Gill, 1993, p. 68)

Knowing can thus be most adequately understood as *the intentional activity of individuals who, as members of a community, make use of and produce representations in the collaborative attempt to better understand and transform their shared world.* (Wells, 1999, p.76)

These conceptions of knowing and learning reflect a shift from knowledge as generalized propositional and symbolic representations internalized by individuals and transferable from context to context—a conception that shaped the development of most current forms of assessment—to knowledge as action, participation, and transformation of individuals within specific social and cultural contexts. Although the most radical constructivists' views are based on clear assumptions, many theorists, such as those just quoted, are not always as explicit about their theoretical assumptions. Most social constructionists referenced here have been heavily influenced by the earlier work of Vygotsky and give a central place to language (more or less explicitly) in understanding learning and how meaning is created in and from spoken and written texts. Wells (1999) explicitly articulates how his work is grounded in both Vygotsky's theory of learning and development and Halliday's language-based theory of learning. Lave and Wenger (1991), on the other hand, make fewer explicit references to the role of language in learning, but it underlies their conception of learning as social practice and coparticipation mediated by language. Learning here is not located in the individual's mind but takes place in interaction with others. It is not always clear how strictly these theorists reject the dualism between the individual mind and the external world, but they all reject the existence of preexisting structures as determining thought and learning.

Because learning cannot be predetermined, learning and teaching are also fundamentally differentiated. In this perspective, “learning can take place where there is teaching but [it] does not take intentional instruction to be itself the source or cause of learning” (Lave & Wenger, 1991, pp. 40–41). Although this is not a novel idea, it is nonetheless a radical departure from the teaching-learning connection implied in movements such as measurement/assessment-driven instruction or standards-based reform, and it has serious implications for educational and assessment practices. A similar distinction concerns the role of theory for educational practices. Lave and Wenger (1991) are quite explicit about this point when they state

undoubtedly, the analytical perspective of legitimate peripheral participation could—we hope that it will—inform educational endeavors by shedding a new light on learning processes, and by drawing attention to key aspects of learning experience that may be overlooked. But this is very different from attributing prescriptive value to the concept of legitimate peripheral participation and from proposing ways of “implementing” or “operationalizing” it for educational purposes. (p. 41)

This perspective is quite different from that of Pellegrino et al., for example, when they state:

the theory-based constructs to be measured must be emphasized prior to test development and then used to generate item and task characteristics that are intended to influence the performance of more or less proficient students. . . . In this way, assessment can be designed with predictable cognitive demands for specific groups of test takers. (pp. 338–339)

The absence of an explicit conceptualization of learning and knowing in most assessment literature seems to be paralleled by an apparent absence of assessment-related issues in the learning literature. Traces of assessment concerns can be found in the work of earlier constructivists (Luria, 1976; Vygotsky, 1935/1978), which led to further work on dynamic assessment (e.g., Brown & Ferrara, 1985; Campione, Brown, Ferrara, & Bryant, 1984; Feuerstein, 1979). In later theoretical formulations (e.g., Lave & Wenger, 1991; Wells, 1999), however, assessment issues are not addressed per se, other than in a few brief statements. In studying the relationship between discourse and knowing, Wells (1999), for example, considers how students’ oral and written responses are generally evaluated as evidence of what they have learned, concluding that “it is important to recognize that no test can tell us what children really think or understand” (p. 106). This might be interpreted as consistent with a constructionist’s view, according to which there cannot be any standards of right and wrong because truth cannot be established with any certainty, and therefore there are no criteria by which

to judge or assess individual construction. But the literature is not clear on this point. Could it be that only a deterministic view of learning allows for a conception of assessment (as we know it now), and that when different epistemological assumptions are made, the very concept of assessment has to be completely reinvented?

To summarize, my purpose here is not to try to resolve the differences between the various theoretical perspectives but to recognize their mutual incompatibility and their implications for the study of learning and for educational practices. In the move from to know, to knowing—that is, from a state of having (knowledge) to an action involving participation, transaction, and transformation—the traditional notion of assessment becomes elusive and the phenomenon under consideration much less tangible. We are moving here from an educational practice of assessment where we have defined a priori what we are looking for, to an educational practice where we are participating in activities in which we formulate representations to better understand and transform the world around us. If our purpose is to understand and support learning and knowing and to make inferences about these phenomena, then it seems that the idea of inquiry—open, critical, and dialogic—rather than that of assessment (as currently understood) would be more helpful, as it would encourage consideration of the epistemological and theoretical assumptions from which we work. This is what I address next by making a case for the usefulness of the inquiry metaphor in learning about learning.

## INQUIRY AND LEARNING

The inquiry literature includes works dealing with various philosophical and theoretical conceptions of inquiry (e.g., Alford, 1998; Blachowicz, 1998; Blaikie, 1993, 2000; Brown, Fauvel, & Finnegan, 1981; Rescher, 1982), with the development of theories of learning and knowing (e.g., Gill, 1993; Wells, 1999), and with the practice of inquiry in teaching and learning (e.g., Bateman, 1990; Brubacher, Case, & Reagan, 1994; Bruce & Bruce, 1992; Short, 1996). There is considerable overlap between the work focusing on learning and that focusing on inquiry. Although these writings reflect different perspectives, audiences, and purposes, the questions they address and the terminology they use are quite similar. In *Dialogic Inquiry*, for example, Wells (1999), concerned with learning and knowing, poses questions about the nature of knowledge, the development and construction of knowing and understanding, and the relationship of knowing to action and inquiry. He states: “it is precisely the process of justifying beliefs through reasoning, conjecturing, evaluating evidence, considering counter-arguments, and so on, that I take to constitute the activity of knowing” (p. 89). He sees knowledge and inquiry as the construction of new understandings and



defines inquiry as “a stance toward experiences and ideas—a willingness to wonder, to ask questions, and to seek to understand by collaborating with others in the attempt to make answers to them” (p. 121). He outlines a conception of education as guided participation in discipline-based inquiry, where understanding of learning<sup>3</sup> develops, and he concludes:

What I am suggesting is that teaching, like learning, is an ongoing process of inquiry, in which the knowledge constructed about the learners and learning, as these are encountered in particular situations, continuously transforms the teacher’s way of understanding and acting in the classroom. (p. 164)

Other writings on inquiry debate philosophical assumptions, the conditions of knowing, the search for explanations and understandings, and the role of theory, methodology, and evidence in developing these understandings. Alford (1998) defines the craft of inquiry as one that “teaches you how to connect theory to evidence in order to construct valid explanations of the workings of society” (p. 18). He sees the construction of these inquiry arguments as an emotional (level of commitment to the task and the capacity to concentrate intensely are crucial) as well as a cognitive process, which includes support from and the participation of colleagues and friends, as well as inspiration from books and from one’s imagination. We can see a parallel here between the inquiry process and the process of learning and developing understanding described by Wells, with its focus on willingness, motivation, and participation with others. Drawing from philosophical hermeneutics for studying the phenomenology of understanding as the aim of qualitative inquiry, Schwandt (1999) conceives of understanding as learning, and as relational in the sense that it is generated from our relationships with others. This argument is helpful for questioning the distinctions often implied in practice between learning, understanding, knowing, inquiring, and so on. How are these different? Learning has traditionally been conceptualized in the context of student learning, whereas inquiry has been more closely associated with social science or scientific inquiry. Might this reflect the hierarchical nature of the activities in which researchers and students are engaged? In other words, experts inquire to generate knowledge (generative activity), and students have to learn or acquire that knowledge (mimetic activity) to become knowledgeable or expert. Or is it that researchers have traditionally engaged in a reflective and detached understanding of the world, making their understandings different from those of students engaged in learning activities, which have traditionally been based on acquiring information produced by others? Are these distinctions still appropriate or relevant given our changing understanding of learning, understanding, and knowing?

What emerges from these writings is an uncanny resemblance between the questions raised about learning and those about inquiry. I am not, however, proposing that these terms be used interchangeably quite yet. Consider this statement: "I am learning about learning." How does it differ from "I am inquiring about learning"? Or if I say "I am inquiring about inquiry," how does it differ from "I am learning about inquiry"? One of the differences that we have established in common usage of these terms (adopted by most current assessment practices) is that learning presumes a particular learned outcome or knowledge (which has most often been definable *a priori*), whereas inquiry refers primarily to an ongoing activity in which one follows certain processes or stages and works from explicit or implicit questions and assumptions about the world and the phenomena being investigated. Also, unlike learning, inquiry does not presume a specific outcome; it is a quest for knowledge and understanding. This distinction, however, is possible only if we conceive of knowledge as fixed and definable, separate from and external to an individual's learning or inquiring. In common usage, *to know* is not an active verb. I cannot say "I am knowing" in the same way I would say "I am learning." We learn and then we know (as in "I know how to ride a bicycle"). Different modes of knowing seem to entail different assumptions about the what and the how of knowing. Drawing from Vygotsky and Luria and a number of scholars working within a cultural-historical tradition, Wells (1999) describes a genesis of knowledge, including six different modes of knowing (i.e., instrumental, procedural, substantive, aesthetic, theoretical, and metaknowing), which, he says, have emerged over long periods of development of human cultures and activities. And although some forms of knowing (e.g., theoretical) have been privileged in certain contexts such as schools, Wells (1999) does not assume the later modes of knowledge to be inherently more valid or useful than earlier modes. These distinctions in modes of knowing, however, might be helpful in sorting out theoretical formulations of learning and in considering the implications for the study of these representations.

Given the absence of theoretical clarity or consensus evident in the literature on learning, knowing, understanding, and so on, it seems that we should remain cautious about the assumptions we make when we engage in the study of their representations, as is being done in current assessment practices. Resisting simplification and stating and questioning the assumptions and perspectives from which we work might be the only way we have to advance our understanding of the phenomena we now claim to assess. We also need, for example, to realize that the language of educational reform we are currently using, such as outcome-based reform or standards-based assessment, is greatly deterministic and places us on epistemological ground that is incompatible with certain other theoretical positions. In effect, current assessment practices and methods prevent us from studying and further understanding the

very phenomena we are claiming to measure. If what we know or understand is inseparable from how we know and understand, how can we justify current assessment practices? I conclude by presenting a set of issues and questions to be considered in an assessment as inquiry perspective.

### CONCLUDING COMMENTS ON ASSESSMENT AS INQUIRY

In previous work (Delandshere, 2001), I argued that current educational assessment is still for the most part a relic of the past: It has limited purposes and methods that generate limited data; it is based on poorly articulated, ad hoc theories and assumptions of learning and knowing; examinees submit to the process without active and equal participation (e.g., critique, reflection, self-reflection); and secrecy, reward, and punishment are still key concepts. Although they are rarely debated in these terms, our assessment practices can be characterized as grounded primarily in empiricist and rationalist epistemological assumptions, where induction and deduction have prevailed as logics of inquiry. But in the last several decades, other forms of social inquiry have developed that seem more consistent with social constructionist (Wittgenstein, 1953) and hermeneutical (e.g., Gadamer, 1989; Heidegger, 1962) assumptions and sociocultural theories of learning. Moss (1996) already argued for including different research traditions in the practice of educational measurement. Thinking of assessment as inquiry indeed opens this inferential process to the broader field of inquiry and its various traditions, which could move current educational and assessment practices beyond the status quo by opening them to scrutiny and debate. Although this will inevitably muddy the water for some time, it will also force us to make explicit the often hidden assumptions from which we work.

In this article, I have analyzed how the very concept of assessment, as we understand it today, exists only within conceptions of knowledge as product of learning, process of learning, or both, definable a priori for categories of individuals and regardless of the broader sociocultural context in which knowledge develops. Given the vastly different theoretical perspectives on learning that are currently at play, how would it even be possible to simply adapt and develop existing assessment methods and analyses when these were conceived within an incompatible epistemological tradition? What does the concept of assessment then become when knowledge is conceived as developing among individuals collectively participating in an activity rather than as some "thing" that individuals possess? The value judgments inherent to assessment certainly do not disappear, but in the latter perspective they are not external to but a part of the act of learning in that they are located in the discourse, actions, and transactions of individuals in participation. How do we understand and document these evaluative judgments,

claims, and responses? Who participates in these evaluative judgments and how? How are these a part of learning? Such inquiry questions are part of understanding and evaluating learning in some sort of meta-learning. As in any inquiry practice, the questions posed, the theories and assumptions involved, the determination of what constitutes data, and the strategies used to generate and interpret data have to be articulated much more purposefully than they typically have been in the context of educational assessment.

If knowing and learning can be conceived of in such different ways, then the aims of these activities and the questions formulated about them are already differentiated. Consider, for example, the typical assessment question "What do students know?" Most current educational assessment practices rest on this question, which is often answered simply by administering tests and scoring the examinees' responses. The test scores are then interpreted as representing the amount or level of knowledge each student possesses with regard to the domain represented by the test. Consistent with early theoretical perspectives on learning (e.g., behaviorism), assessment questions have typically been defined in narrow and oversimplified ways (e.g., "How is Jenny doing in spelling this grading period?"). From a different theoretical perspective (e.g., sociocultural), however, questions such as "What do students know?" or even "How do they accomplish this task?" could not be posed or answered so simply. If what we know cannot be separated from how we know and from the experiences and activities that shape it, then the assessment questions have to be framed in such a way as to be consistent with this theoretical perspective and include social, cultural, and ethical issues that have not been typically addressed. The questions asked would therefore be different. For example, accountability questions could not be answered merely by measuring predetermined learning outcomes of individual students. Instead of measuring what students do not know with regard to a fixed domain, students' educational experiences and the activities in which they engage would need to be considered to understand what and how they know and learn. If learning is a "kind of doing" (Gill, 1993, p. 68), or the transformation and change of people (Lave & Wenger, 1991), then in documenting learning we have to take its dynamic nature into account. In this perspective, because individuals' experiences will always be different, the current focus on the assessment of individual differences in learning outcomes at a given moment and regardless of the context and experiences of these individuals becomes irrelevant and meaningless.

Who defines the assessment/inquiry questions and who participates in the inquiry is another issue that needs to be addressed. Current educational assessment practices implicitly assume an authoritative relationship between the assessor (the knower) and the assessee (the learner). The evaluative judgments result from independent and objective observations of the latter by the former, leading to assessments that are static and unidirec-

tional, hence reproducing forms of knowledge and maintaining the status quo. Is this duality assumption universal and consistent with all the theoretical perspectives considered here? Can we imagine evaluative judgments of learning that are more dialogic and that result from the interactions among participants in the activity of learning? What becomes of external assessment (such as current forms of accountability testing) imposed from outside the activity of learning, which seems to be a product only of deterministic theoretical conceptions of learning?

The concern with uncovering theories and assumptions is not simply an academic exercise but rather the only way we have to understand our practices and reformulate the questions and perspectives from which we work. Theories can be formulated at many different levels, from broad theories of social action, of human experiences, to more specific theories of learning and cognition within particular disciplines, for example. One encompasses the others as it accounts for broader social phenomena, but the connections between the related theoretical formulations are not always made in practice. Educational measurement practices have not been explicitly grounded in social action theories or in learning theories, for that matter. Cognitive theories are sometimes limited to a set of task descriptions. But when assumptions and theoretical explanations are not made explicit, they tend to appear to be unreasoned speculations. So, for example, when the same test is given to all sixth graders in a state to find out whether their educational experiences yield similar achievements, is it because we are working from a theory stating that if students have all been taught the same thing, they all will learn it in the same way at the same time? It seems unlikely that any educator would articulate such a theory. Yet without this perspective, how can current forms of state-mandated assessment be justified? In this practice a great disparity seems to exist between the educational and political assumptions.

The determination of what constitutes data and the strategies used to generate and interpret the data are an unresolved issue. The incompatibility between new theories of learning and the methods currently used in assessment has been largely recognized (e.g., Embretson, 1993; Haertel & Wiley, 1993; Mislevy et al., 2001; Pellegrino et al., 1999). In educational assessment, data-collection methods and analysis are typically considered to be a technical matter (e.g., scoring accuracy, probabilistic models, consistency of responses) and are selected with a concern for efficiency (Madaus & O'Dwyer, 1999). As different theoretical conceptions of learning and knowing call on different inquiry traditions, the ways of knowing also change, along with the methods and procedures used. What constitutes data and evidence also has to be redefined. If other, more dialogic and deliberative means of developing understandings are considered, the process becomes less technical, more interpretive, and more consistent with these theoretical perspec-

tives. The main point here is to establish the link between the inquiry/assessment questions, the theoretical perspectives, and the strategies used to construct answers to the questions. Theories and assumptions remain invisible unless rationales and evidence are provided for the claims we make and the actions we take. Only when the links between theoretical claims, arguments, and practices are articulated—hence opened for scrutiny and debate—will change in educational practices be possible. Furthermore, if this inquiry into learning is to further the understanding of learning, it will be critical to pose questions and consider data that will allow us to challenge and reconstruct the theories and assumptions from which we work.

I do not intend here to crystallize or unnecessarily formalize the process of inquiry into learning (or the assessment process) but rather to open a dialogue and encourage greater participation in the decisions and judgments that are made in and about our educational practices. These considerations may not all be relevant to all cases of learning, but such conclusions should be the result of careful consideration rather than based on implicit assumptions and unexamined theories. It is also possible that in some cases these considerations will need to be redefined or new ones formulated. In other words, new cases will emerge, and new questions will be raised. Thus, any frameworks or considerations proposed to study and assess learning will need to evolve and change because these remain useful only to the extent that they help us better understand our practice. The intent is not to propose a solution here but simply to offer a way to think differently about our educational practices. These concluding comments are purposely left open because it is difficult to envision a practice of inquiry shorn of the implicit assumptions about learning that have remained deeply entrenched for more than a century. Clearly, when we attempt to rethink one aspect of our educational practice, the whole enterprise comes into question. It is not possible to rethink educational assessment while leaving other educational practices intact and vice versa. Working from broader theories of social action has enormous implications for how we might organize learning, teaching, and schooling.

*A version of this article was presented at the annual meeting of the American Educational Research Association, New Orleans, April 2000. I would like to thank Gretchen Guiton for the ongoing and insightful conversation we have had about assessment, Terrence Mason for reading and commenting on multiple drafts of this article, and the anonymous reviewers for their helpful suggestions.*

## Notes

1 As classroom assessment practice tends to emulate mandated assessment, it does not seem sensible to address one without the other. It also does not seem useful to define assessment differently in these two contexts.

2 Term borrowed from D. C. Phillips.

3 One could also say the understanding of understanding; others might refer to this idea as assessment.

## References

- Alford, R. R. (1998). *The craft of inquiry: Theories, methods, evidence*. New York: Oxford University Press.
- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks-Cole.
- Bateman, W. L. (1990). *Open to question: The art of teaching and learning by inquiry*. San Francisco: Jossey-Bass.
- Blachowicz, J. (1998). *Of two minds: The nature of inquiry*. Albany: State University of New York Press.
- Blaikie, N. (1993). *Approaches to social inquiry*. Cambridge, MA: Blackwell.
- Blaikie, N. (2000). *Designing social research*. Malden, MA: Blackwell.
- Brown, A., & Ferrara, R. (1985). Diagnosing zones of proximal development. In J. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives*. New York: Cambridge University Press.
- Brown, S., Fauvel, J., & Finnegan, R. (1981). *Conceptions of inquiry*. New York: The Open University Press.
- Brubacher, J., Case, C., & Reagan, T. (1994). *Becoming a reflective educator: How to build a culture of inquiry in the schools*. Thousand Oaks, CA: Corwin.
- Bruce, W., & Bruce, J. (1992). *Learning social studies through discrepant event inquiry*. Annapolis, MD: Alpha Publishing Company.
- Campione, J., Brown, A., Ferrara, R., & Bryant, R. (1984). The zone of proximal development: Implications for individual differences and learning. In B. Rogoff & J. Wertsch (Eds.), *Children's learning in the zone of proximal development*. San Francisco: Jossey-Bass.
- Cannella, G. S. (1998). Early childhood education: A call for the construction of revolutionary images. In W. E. Pinar (Ed.), *Curriculum: Toward new identities*. New York: Garland.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302.
- Delandshere, G. (2001). Implicit theories, unexamined assumptions and the status quo of educational assessment. *Assessment in Education: Principles, Policy and Practice*, 8(2), 113–133.
- Delandshere, G., & Jones, J. H. (1999). Elementary teachers' beliefs about assessment in mathematics: A case of assessment paralysis. *Journal of Curriculum and Supervision*, 14(3), 216–240.
- Delandshere, G., & Petrosky, A. (1998). Assessment of complex performances: Limitations of key measurement assumptions. *Educational Researcher*, 27(2), 14–24.
- Douglas, J. D. (1971). *Understanding everyday life*. London: Routledge and Kegan Paul.
- Embretson, S. E. (1993). Psychometric models for learning and cognitive processes. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- Embretson, S. E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods*, 3, 380–396.
- Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1993). Learning, teaching, and testing for complex conceptual understanding. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- Feuerstein, R. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. Baltimore: University Park Press.
- Feyerabend, P. K. (1978). *Against method: Outline of an anarchistic theory of knowledge*. London: Verso.
- Frederiksen, N., Mislevy, R. J., & Bejar, I. I. (1993). *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.



- Gadamer, H. G. (1989). *Truth and method* (2nd ed.). New York: Crossroad Publishing.
- Gill, J. H. (1993). *Learning to learn: Toward a philosophy of education*. Atlantic Highlands, NJ: Humanities Press.
- Gipps, C. V. (1994). *Beyond testing*. London: Falmer.
- Gipps, C. V. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24.
- Glaser, R. (1976). Components of a psychology of instruction: Toward a science of design. *Review of Education Research*, 46, 1–24.
- Haertel, E. H., & Wiley, D. E. (1993). Representations of ability structures: Implications for testing. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- Heidegger, M. (1962). *Being and time*. New York: Harper & Row.
- Hopkins, K. C., Stanley, J. C., & Hopkins, B. R. (1990). *Educational and psychological measurement and evaluation*. New York: Prentice Hall.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–230). Cambridge, UK: Cambridge University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Lohman, D. F., & Ippel, M. J. (1993). Cognitive diagnosis: From statistically based assessment toward theory-based assessment. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.
- Madaus, G. F. (1994). Testing place in society: An essay review of testing—Social consequences of examined life. *American Journal of Education*, 102, 222–234.
- Madaus, G. F., & McDonagh, J. T. (1979). Minimum competency testing: Unexamined assumptions and unexplored negative outcomes. *New Directions for Testing and Measurement*, 3, 1–15.
- Madaus, G. F., & O'Dwyer, L. M. (1999). A short history of performance assessment. *Phi Delta Kappan*, 688–695.
- Masters, G. N., & Mislevy, R. (1993). New views of student learning: Implications for educational measurement. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- McCarty, L., & Schwandt, T. A. (2000). Seductive illusions: Von Glasersfeld and Gergen on epistemology and education. In D. C. Phillips (Ed.), *Constructivism in education*. NSSE. Chicago: University of Chicago Press.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13–23.
- Mislevy, R. J. (1993). Foundations of a new test theory. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- Mislevy, R. J., Steinberg, L. S., Almond, R. G., Haertel, G. D., & Penuel, W. R. (2001). *Leverage point for improving educational assessment* [CSE Technical Report 534]. Los Angeles: Center for the Study of Evaluation and National Center for Research on Evaluation, Standards and Student Testing, University of California, Los Angeles.
- Moss, P. A. (1996). Enlarging the dialogue in educational measurement: Voices from interpretive research traditions. *Educational Researcher*, 25(1), 20–28, 43.
- Pellegrino, J. W., Baxter, G. P., & Glaser, R. (1999). Addressing the “Two Disciplines” problem: Linking theories of cognition and learning with assessment and instructional practices. *Review of Research in Education*, 24.
- Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher*, 24(7), 5–12.

- Phillips, D. C. (2000). An opinionated account of the constructivist landscape. In D. C. Phillips (Ed.), *Constructivism in education*. Chicago: University of Chicago Press.
- Polanyi, M. (1958). *Person knowledge: Towards a post-critical philosophy*. London: Routledge and Kegan Paul.
- Popham, W. J. (1990). *Modern educational measurement: A practitioner's perspective* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Rescher, N. (1982). *Empirical inquiry*. Totowa, NJ: Rowman & Littlefield.
- Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. Gifford & M. O'Connor (Eds.), *Changing assessment: Alternative views of aptitude, achievement and instruction* (pp. 37–76). London: Kluwer Academic Publishers.
- Schwandt, T. (1999). On understanding understanding. *Qualitative Inquiry*, 5(4), 451–464.
- Shepard, L. A. (1989, April). Why we need better assessment. *Educational Leadership*.
- Shepard, L. A. (1991). Psychometricians' beliefs about learning influence testing. *Educational Researcher*, 20(7), 2–16.
- Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 2–14.
- Short, K. (1996). *Learning together through inquiry*. York, ME: Stenhouse.
- Skinner, B. F. (1954). The sciences of learning and the art of teaching. *Harvard Educational Review*, 24, 86–97.
- Slife, B. D., & Williams, R. N. (1995). *What's behind the research?* Thousand Oaks, CA: Sage.
- Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for education measurement. In R. Linn (Ed.), *Educational Measurement* (3rd ed.; pp. 263–331). New York: Macmillan.
- Snow, R. E., & Lohman, D. F. (1993). Cognitive psychology, new test design, and new test theory: An introduction. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- Solso, R. L. (1991). *Cognitive psychology*. New York: Allyn & Bacon.
- Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher*, 27(3), 11–20.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press. (Original work published 1935)
- Wallace, W. L. (1971). *The logic of scientific sociology*. Chicago: Aldine.
- Wells, G. (1999). *Dialogic Inquiry: Towards a sociocultural practice and theory of education*. New York: Cambridge University Press.
- Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco: Jossey-Bass.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford, UK: Blackwell.
- Zimiles, H. (2000). On reassessing the relevance of the child development knowledge base to education. *Human Development*, 43, 235–245.

GINETTE DELANDSHERE is an associate professor in the School of Education at Indiana University in Bloomington. Her work focuses on the study and teaching of inquiry methodology and the assessment of teaching and learning. She is the author of "Implicit Theories: Unexamined Assumptions and the Status Quo of Educational Assessment," *Assessment in Education: Principles, Policy and Practice*, 8(2), 113–133 and coauthor, with Sheila Arens, of "Examining the Quality of the Evidence in Pre-Service Teacher Portfolios," *Journal of Teacher Education* (2003).